

6D-01

エッジコンピューティングによる 映像ストリームの効率的な中間処理方法の検討

藤野知之¹ 柏木啓一郎¹

¹NTT ソフトウェアイノベーションセンタ

1 背景と課題

Web カメラ等の IoT 機器の普及や市街のスマートシティ化により様々な場面で映像が蓄積・利用されている。多数の映像データを有効活用しようという試みは、スマートシティにおける信号の最適化のための交通量監視 [1] や市中の様々な異常検知 [2] など多数見られる。深層学習による画像処理技術の進展により、映像の情報を有効活用する方法も確立されてきており、従来難しかったようなサービスが実現可能になっている。

一方で、それらの映像には個人情報に類する情報も含まれており、プライバシー権の保護の観点から取り扱いに注意が必要な情報にマスク処理を行ったのち、データを蓄積するという事が行われている [3]。例えば、人の顔や車のナンバープレートにマスクをかけたり、Web 会議において室内にマスクをかけたり、といった処理が広く行われている。

本研究では、カメラと映像を保存するデータストアの中間においてこのマスク処理を行う、「エッジコンピューティング」により、データが永続的に保存される前段階でマスクングを施し、個人情報の流通範囲を最小限に抑えることを考える。

カメラの数が増えると、このマスクングの処理に必要なコストの増大が課題となる。特に映像の符号化・復号に伴う処理は処理量が多い。図 1 上は設置カメラから得られる映像ストリームにマスクング処理を施すパイプラインの例である。映像データは画像（フレーム）の集合であるが、フレーム内や、前後のフレーム間の冗長性を排するような符号化による圧縮が施されているため、マスクングを行うためには一度画像に復号したのちマスク処理を施し再度符号化するのが一般的である。

本研究では、この復号と再符号化のコストの削減を図るため、図 1 下の提案パイプラインのように最低限の情報のみ復号し、マスク位置を割り出した後、映像符号に

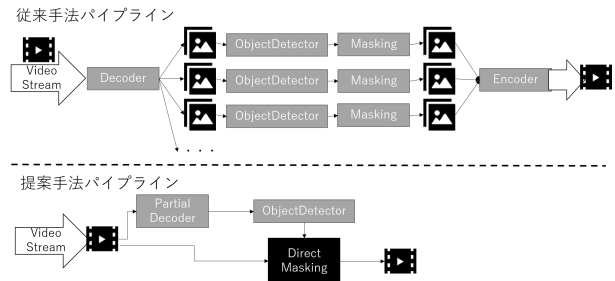


図 1 映像ストリームのマスクング処理パイプライン

直接マスクを書き込む方法を検討している。本稿では、このパイプラインのうち、映像の符号の一部を直接上書きするマスクング法に着目し、実験により実現可能性の検証を行った。

2 提案方法：H.264 映像符号への直接マスク書き込み

本稿では映像符号化方式として市中で最もよく使われているものの一つである H.264 形式を取り上げる。H.264 には様々なオプションが存在するが本稿ではもっとも基本的な制約付きベースラインプロファイルを用いる。

符号化された映像データは画像データを内包する I フレームと動きを補償する P フレームに大別することができる。図 2 に示すように、I フレームによってまず画像が描画されたのち、P フレームはその前段の I フレームを参照することで I フレームからの動きを表現する。画像が主に I フレームに含まれるため、マスクを施すためには I フレームの一部を操作する必要がある。

各フレームはそれぞれヘッダ情報とマクロブロックという縦横各 16pixel の小画像の連続で構成される。本稿では最も単純な直接マスク書き込み方法として、マクロブロック境界を取得し、マスク対象のマクロブロックを空白のマクロブロックと置き換える実装を行った。

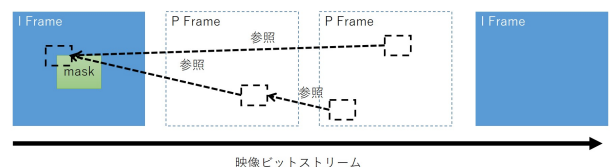


図 2 符号化された映像のストリーム

* A Study for Efficient Movie Processing by Using Edge Computing
Tomoyuki Fujino and Keiichiro Kashiwagi
NTT Software Innovation Center

3 実装と実験

実装のベースとしてオープンソースのソフトウェアデコーダー・エンコーダーである OpenH264 を改変した。マクロブロックは CAVLC(Context Adaptive Variable Length Coding) という符号で保存されており、符号が可変長であるためマクロブロックの境界を見つけるためにはフレームを先頭からパースする必要がある。ただし、CAVLC 符号を画像に再構成する必要はないため通常の復号処理よりも軽量である。

境界を見つけたのち、マスク対象のマクロブロックを空白のマクロブロックで上書きを行った。空白マクロブロックによる機械的に上書きを行えるように、以下のような固定値のパラメータを指定した。

表 1 空白マクロブロック設定値

パラメータ	設定値
MB TYPE	INTRA 16x16
Luma 16x16 Mode	I16 PRED DC
CBP	0
DeltaQP	0
nC/Total Coefficients/Trailing Ones	0/0/0

3.1 実験

図 3 の Web カメラによる映像のうち、図中の枠線部分に相当するマクロブロックに対して提案方法による直接マスク書き込みを行った。その後復号した結果が図 4 である。マスクした場所はモザイク状になり、情報が失われていることが確認できた。しかしながら、画面全体の色味が崩れ、ブロックノイズのような現象が生じてしまった。また、マスク位置によっては映像が再生できなくなってしまう場合があった。

MacBook Pro 2018 年モデル (CPU: Intel Core i9-8950HK) 上で動作検証を行ったところ、約 5 秒間の Web カメラ映像に対して、オープンソースのソフトウェアエンコーダーである FFMPEG を用いて従来パイプラインによるマスク処理を施した場合 2.643 秒かかった。一方、本提案方法を用いたときの処理時間は 0.744 秒であり、本提案方法による高速化の可能性を示した。

4 考察と今後の課題

実験により効率よいマスクングを行える可能性が示された。しかしながら、画面が崩れてしまう問題が生じた。これは、本実装がマクロブロックの上書きをする際、表

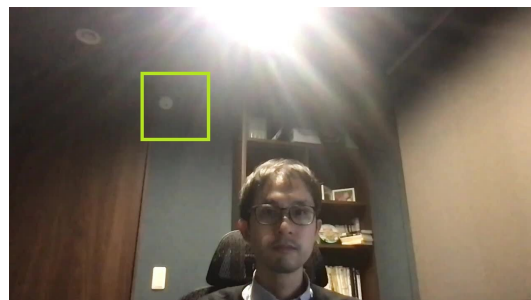


図 3 元映像の I フレームの復号結果



図 4 提案方法によりマスクした I フレームの復号結果

1 のように定値で書き込んでいることに起因すると考えられる。マクロブロックは独立ではなくパラメータが周囲のマクロブロックとの依存関係を持ち、周囲のマクロブロックと連動していない定値書き込みがノイズとなり、画面全体に伝搬したのと考えられる。

また、マスク位置によっては再生不可になってしまった。これも画面崩れ同様の問題が生じ、ノイズにより画面内のマクロブロックが指定可能範囲外のパラメータを持ってしまったものと思われる。これらの課題は、マクロブロックの設定値を適切に設定することで影響を軽減することができると思われる。

今後は画面崩れを防ぎ、実用可能な品質にするとともに、H.264 の他のプロファイルや AV1 などさらに最新のコーデックで対応できるか検討を進めていきたい。

参考文献

- [1] Dinesh Singh, C Vishnu, and C Krishna Mohan. Visual big data analytics for traffic monitoring in smart city. In *2016 15th IEEE international conference on machine learning and applications (ICMLA)*, pages 886–891. IEEE, 2016.
- [2] Sharmila Chackravathy, Steven Schmitt, and Li Yang. Intelligent crime anomaly detection in smart cities using deep learning. In *2018 IEEE 4th International Conference on Collaboration and Internet Computing (CIC)*, pages 399–404. IEEE, 2018.
- [3] Alem Fitwi, Yu Chen, Sencun Zhu, Erik Blasch, and Genshe Chen. Privacy-preserving surveillance as an edge service based on lightweight video protection schemes using face de-identification and window masking. *Electronics*, 10(3):236, 2021.